WE CLAIM:

1. A method of predicting the occurrence of a selected weather event, the method comprising the steps of:

identifying a first set of parameters that relate to the selected weather event, wherein the parameters in the first set are statistically correlated;

identifying a data set comprised of a plurality of observations for the weather event over a selected geographical area, each observation containing values for the parameters;

performing a Principal Component Analysis (PCA) to identify and select a reduced set of principal components corresponding to a second set of parameters, the reduced set of principal components being uncorrelated, the second set of parameters being a subset of the first set of parameters, the parameters in the second set of parameters contributing at least a majority of the variance of the first set of parameters;

rotating the reduced set of principal components to provide a third set of parameters, wherein each parameter corresponds to a physical data measurement; and

performing a logistic regression on the third set of parameters to determine the predictive equation for forecasting of the selected weather event.

2. The method of Claim 1, wherein the step of rotating the second set of parameters is accomplished by a linear transformation.

3. The method of Claim 1, wherein the step of identifying the data set of observations comprises the formation of a data matrix $[X]$ having n rows and m columns, each row representing the i-th observation and each column representing the j-th parameter, where i is an index from 1 to n and j is an index from 1 to m.

4. The method of Claim 3, wherein the step of performing the Principal Component Analysis (PCA) further comprises the step of

calculating a scaled data matrix $[Z]$ according to the following equation:

$$[Z]=[D]^{-1}\left([X]-\frac{1}{n}[1][X]\right)$$

where $\frac{1}{n}[1][X]$ is the mean of $[X]$, $[1]$ is the unity matrix whose elements are all equal to unity, and $[D]^{-1}$ is the diagonal matrix with diagonal elements being the reciprocals of the sample standard deviations of the *m* parameters and the remaining element being 0,
whereby scaling the data allows parameters with different magnitudes of variance to be equally considered.

5. The method of Claim 4, wherein the step of performing the Principal Component Analysis (PCA) further comprises the step of

calculating a correlation matrix $[R]$ from the scaled data matrix $[Z]$ according to the following equation:

$$[R]=\frac{1}{n-1}[Z]^T[Z]$$

6. The method of Claim 5, wherein the step of performing the Principal Component Analysis (PCA) further comprises the step of

calculating an eigenvector matrix $[E]$ and an eigenvalue matrix $[\Lambda]$, both of which are derived from the correlation matrix $[R]$.

7. The method of Claim 6, wherein the step of performing the Principal Component Analysis (PCA) further comprises the step of

calculating an PC loading matrix $[A]$ according to the following equation:

$$[A]=[E][\Lambda]^{1/2}$$

8. The method of Claim 7, wherein the step of performing the Principal Component Analysis (PCA) further comprises the step of

ordering the parameters by descending values of the eigenvalues calculated for each parameter; and

selecting the p parameters that account for at least a majority of the variance observed in the observations.

9. The method of Claim 8, wherein selecting the p parameters is accomplished by plotting the parameters in descending order on a graph and selecting those parameters that are to the left of where the slope of the graph of the parameters approaches a selected distance from 0.

10. The method of Claim 1, wherein the definition of the geographical area comprises a regular grid.

11. The method of Claim 1, wherein the definition of the geographical area comprises a set of discrete points.

12. The method of Claim 1, wherein the predictive equation for forecasting of the selected weather event provides a probability of a binary event.

13. A system for forecasting the probability of occurrence of a selected weather event, the system comprising:

selecting a geographical area, the geographical area being divided into a regular grid;

developing predictors for observed values of atmospheric phenomena within the geographical area;

collecting a plurality of observations of the predictors for each portion of the grid over a period of time;

forming a data matrix $[X]$ having *n* rows and *m* columns, each row representing the *i*-th observation and each column representing the *j*-th parameter, where *i* is an index from 1 to *n* and *j* is an index from 1 to *m*;

normalizing the data within the data matrix to provide a scaled data matrix $[Z]$, wherein parameters having different magnitudes of variance may be equally weighted;

calculating a correlation matrix $[R]$ from the scaled data matrix $[Z]$;

calculating an eigenvector matrix $[E]$ containing eigenvectors calculated from the correlation matrix $[R]$ and an eigenvalue matrix $[\Lambda]$ containing eigenvalues calculated from the respective eigenvectors, wherein the eigenvector matrix $[E]$ provides a set of orthogonal vectors relating to variance;

calculating a PC loadings matrix $[A]$ from the eigenvector matrix $[E]$ and the eigenvalue matrix $[\Lambda]$;

calculating the PC score matrix $[F]$ from the scaled data matrix $[Z]$ and the PC loadings matrix $[A]$; and,

performing a logistic regression using the PC score matrix $[F]$ to provide a probability for the occurrence of the selected weather-related event.

14. The system of Claim 13, further comprising:

forming a reduced set of eigenvectors, wherein the eigenvectors in the reduced set correspond to parameters that account for at least a majority of the total variance of the parameters in the data set;

performing an orthogonal linear transformation on the reduced set of eigenvectors for form a rotated set of eigenvectors; and,

calculating the PC loadings matrix $[A]$ calculated from an eigenvector matrix $[E]$ consisting of the rotated set of eigenvectors and an eigenvalue matrix $[\Lambda]$ consisting of the eigenvalues corresponding to the rotated set of eigenvectors.

16. A computer program product for forecasting the occurrence of a weather event, the computer program product embodied on one or more computer-readable media and comprising:

computer-readable program code means for input of a data set comprised of observations of data parameters made during a known time period at geographical sites;

computer-readable program code means for performing a Principal Component Analysis (PCA) to identify and select a reduced set of principal components corresponding to a subset of the data parameters, the reduced set of principal components being uncorrelated, the parameters in the subset of parameters contributing at least a majority of the variance of all the parameters;

computer-readable program code means for performing a logistic regression to determine the predictive equations for forecasting the occurrence of the selected weather event.

16. The computer program product described in Claim 15, wherein the computer-readable program code means for performing a Principal Component Analysis (PCA) comprises

a means for forming a data matrix $[X]$ having n rows and m columns, each row representing the i-th observation and each column representing the j-th parameter, where i is an index from 1 to n and j is an index from 1 to m;

a means for normalizing the data within the data matrix to provide a scaled data matrix $[Z]$, wherein parameters having different magnitudes of variance may be equally weighted;

a means for calculating a correlation matrix $[R]$ from the scaled data matrix $[Z]$;

a means for calculating an eigenvector matrix $[E]$ containing a first set of eigenvectors calculated from the correlation matrix $[R]$ and an first eigenvalue matrix $[\Lambda]$ containing eigenvalues calculated from the respective eigenvectors, wherein the eigenvector matrix $[E]$ provides a set of orthogonal vectors relating to variance;

a means for identifying an eigenvector matrix $[E]$ consisting of a second set of eigenvectors corresponding to parameters that account for at least a majority of the total variance of the parameters in the data set;

a means for performing an orthogonal linear transformation on the second set of eigenvectors for form a rotated set of eigenvectors; and,

a means for calculating a PC loadings matrix $[A]$ from the eigenvector matrix $[E]$ consisting of the rotated set of eigenvectors and the eigenvalue matrix $[\Lambda]$ consisting of the eigenvalues corresponding to the rotated set of eigenvectors;

a means for calculating a PC loadings matrix $[A]$ from the eigenvector matrix $[E]$ and the eigenvalue matrix $[\Lambda]$ corresponding to the rotated set of eigenvectors; and,

a means for calculating the PC score matrix $[F]$ from the scaled data matrix $[Z]$ and the PC loadings matrix $[A]$, the PC score matrix $[F]$ provided as input to the logistic regression.